15B.5

RAINFALL ATTRACTORS AND PREDICTABILITY

Aitor Atencia *, Isztar Zawadzki and Frédéric Fabry
McGill University, Montreal, QC, Canada

1. ABSTRACT

The atmosphere has been related to chaotic systems ever since Eduard Lorenz's influential paper of 1963. However, determining the existence of an atmospheric attractor and its intrinsic predictability is still an unresolved problem.

In this study, 15 years of US composite radar is analyzed in an attempt to shed light on these questions. First, the atmospheric or rainfall field's attractor is examined in a three-dimensional phase space defined by three weakly correlated properties of the rainfall fields. The fractal properties of this attractor are studied and compared to other well-known systems such as the Lorenz attractor to determine scaling processes within rainfall fields.

Then, the intrinsic predictability of rainfall fields is mapped in the previously defined phase space. This predictability map shows a structure that reveals how the initial statistical properties of a rainfall field are related to its predictability. Consequently, this information would allow us to assess forecast quality of future events using only the initial conditions.

This work also proofs statistically some well-established beliefs about predictability of rainfall systems, such as the high predictability of frontal rainfall systems and the low predictability of isolated convective storms.

2. INTRODUCTION

While there is an agreement of using ensemble prediction systems of perturbed initial condition (Toth and Kalnay, 1993) for predicting the weather, optimal methods to predict the predictability are still in debate (Smith et al, 1999).

Since the acceptance that the atmosphere is a chaotic system (Lorenz, 1963), different studies have looked for attractors in worldwide climate (Grasberger, 1986) or temperature (Nicolis and Nicolis, 1984) but without investigating the predictability consequences of the existence of this attractor.

On the other hand, some researchers have studied the intrinsic predictability from ensemble NWP systems (Melhauser and Zhang, 2012) or radar images (Carbone et al, 2002). Even some studies have focused on predictability of precipitation from convective adjustment time-scale (Keil, 2013) or large-scale forcing (Jankov and Gallus, 2004).

In this study a different approach is taken to assess the predictability by revisiting the chaos theory.

* *Corresponding author address:* Aitor Atencia, McGill University, Dept. of Atmospheric and Oceanic Sciences, Montreal, QC, Canada
e-mail: aitorrent@gmail.com

3. DATA

A large dataset is required to establish statistically robust conclusions about attractors and predictability. With this purpose, two composite mosaics over North America have been used. The first one is the NOWrad mosaic produced by the Weather Services International (WSI) Corporation. NOWrad is a three-step quality controlled product with a 15-minute temporal resolution and 2 km spatial resolution. These mosaics show the maximum reflectivity measured by any radar at each grid point at any of the 16 vertical levels. This data is available for the period from October 1995 to December 2007. The second dataset is produced by Weather Decision Technologies (WDT) and uses radar data from the entire Weather Surveillance Radar-88 Doppler (WSR-88D) network in the Continental United States (CONUS). This allows WDT to apply their most up-to-date, technologically advanced algorithms to provide superior quality radar data through the removal of false echoes and through the blending of multiple radars. WDT creates seamless radar mosaics with a high spatial resolution (1 km) and temporal resolution (5 min) from January 2004 to April 2011.

A common grid is defined for the two data set used in this study (Fig. 1). The new 512x512 point's grid has a 4 km resolution. Each reflectivity map, Z, is converted to rainfall rate, R, by the relation, $Z=300 \bullet R^{1.5}$. The rainfall values are interpolated to the new grid and the highest value from both sources is chosen for the common period (01/2004 to 12/2007). After the up-scaling process in rainfall units the obtained field is converted to reflectivity. The temporal resolution of the new data set is 15 minutes. The selected domain avoids the Rocky Mountains because their orographic effects over rainfall fields and the blockage they produce in the radar rainfall images.

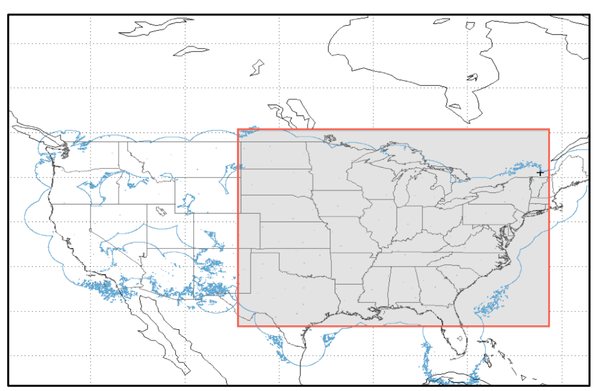

Figure 1.- Location domain. The red rectangle corresponds to the common domain on which all the reflectivity fields are smoothed. The blue contours represent the coverage of the reflectivity mosaics.

4. PREDICTABILITY IN LORENZ SYSTEM

In 1963, Lorenz developed a three dimensional simplified mathematical model for atmospheric convection. His mathematical model, also known as Lorenz system, consists of three nonlinear ordinary differential equations representing the phase space evolution and it is formulated as:

$$\frac{dx}{dt} = \sigma(y - x)$$
$$\frac{dy}{dt} = x(\rho - z) - y$$
$$\frac{dz}{dt} = xy - \beta z$$

Here *x*, *y* and *z* make up the system state, *t* is time, and *σ*, *ρ*, *β* are the system parameters.

The Lorenz original parameter choice was *σ=10*, *ρ=28*, *β=8/3*. For these parameters the system is drawn towards a strange attractor (Fig. 2). An attractor is a set towards which a variable, moving according to the dictates of a dynamical system, evolves over time. An attractor is called strange when it has a fractal structure. A fractal is an object with some degree of self-similarity (exact or statistical) and with structure at arbitrarily small scales. Self-similarity means that object looks the same irrespective of the scales at which you inspect it. The similarity dimension (or box-counting dimension) is the simplest way to measure the degree of self-similarity. It is defined as the number of boxes *N(ε)* required to cover the object scales with the size (edge-length, *ε*) of the boxes. The dimension (*d*) is formulated as:

$$N(\varepsilon) \propto 1/(\varepsilon^d)$$

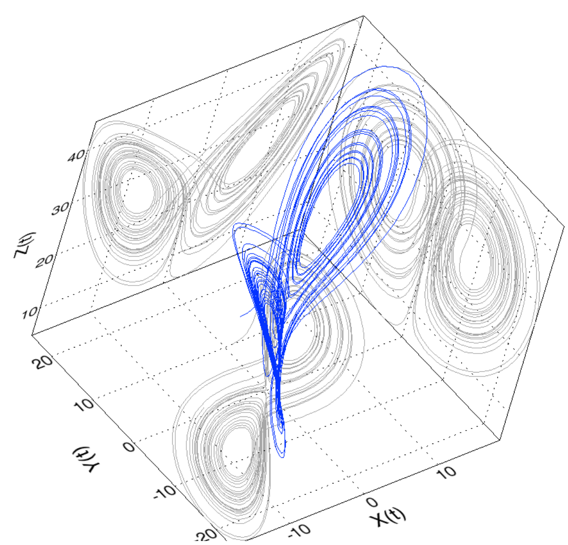

Figure 2.-3D plot of the Lorenz attractor in the phase space and 2D projection in the phase planes.

The calculation of this dimension in practice requires a prohibitive amount of time for experimental dynamical system. Therefore, the most widely used dimension estimation is the correlation dimension (*Dc*). To define *Dc*, the correlation sum, Cr, is defined as:

$$C_r = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1; j\neq i}^{N} \Theta\left(r - |X_i - X_j|\right)$$

Where *Θ* is the Heaviside function, *r* is the radius of a sphere centered on a phase space point, X_i, on the attractor trajectory. The total number of points in the phase space is *N*.

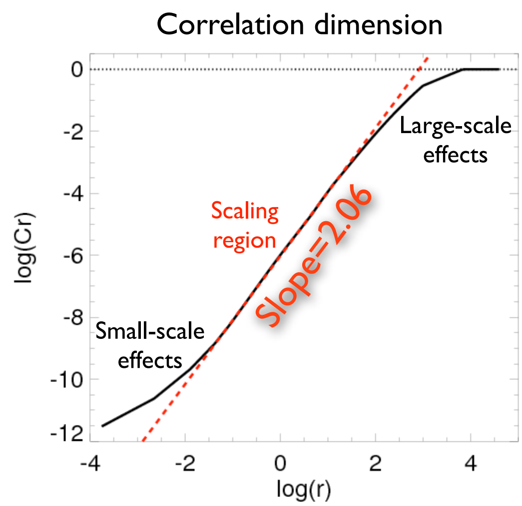

Figure 3.-The log(r)-log(Cr) plot for determining the correlation dimension by means of the scaling region slope.

The correlation sum scales with the radius (Fig. 3) following a power law relation.

$$C_r \ \alpha \ r^{Dc}$$

Where the exponent, *Dc*, is the correlation dimension.

The Lorenz attractor (Fig. 2) assumes the famous “butterfly wing” pattern in the phase space. Applying the box-counting to the strange attractor of Lorenz model this gives *d* = 2.05. The correlation dimension technique (Fig. 3) gives *Dc* = 2.06. Fractals usually have non-integer dimensions. Strange attractors appear in chaotic system.

Consequently, Lorenz system is a chaotic system, which means it is impossible to have long-term prediction. This is because these systems are highly sensitive to initial conditions. Small differences in initial conditions (such as those due to rounding errors in numerical computation) yield widely diverging outcomes for such dynamical systems, rendering long-term prediction impossible in general.

Predictability is the degree to which a correct prediction or forecast of a system's state can be made. In chaotic systems, different measurements can be used to measure the predictability of an initial point:

1. Doubling times: $T_d = \log(2)/\log(1 + r/100)$
2. Lead-time (T_l): It's defined as the saturation time of the error.
3. Tangent-linear time: $r = \mathrm{Cov}(\delta^+, \delta^-)/\sqrt{Var(\delta^+, \delta^-)}$; When correlation has 0.5 value that corresponds to random perturbations, so Tangent-linear time (T_t) is defined at this point.

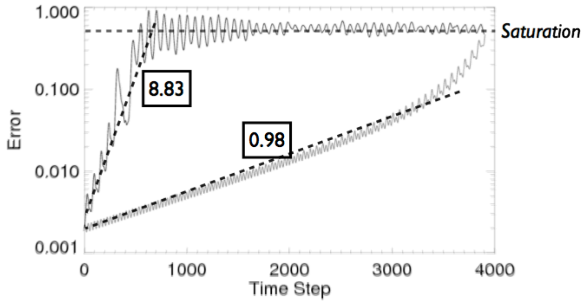

Figure 3.-RMSE error evolution in function of lead-time for two different Initial points of the Lorenz system.

Figure 4 shows the error (defined as the root mean square difference between the three phase state variables) in function of lead-time for two different Initial conditions.

Computing the three prediction times previously defined for different initial points in the space of (x,y,z) ∍ [-20,20 ; -20,20 ; 0,40], it can be observed (Fig. 5) the three indexes provides the same information and, consequently, they are equivalent.

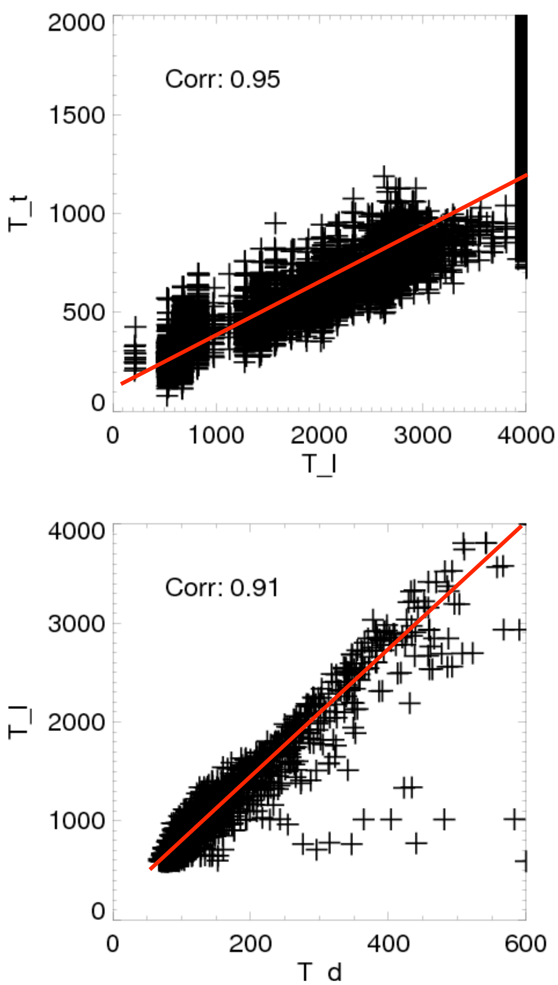

Figure 2.-Correlation between tangent-linear time and Lead-time (upper figure) and between lead-time and doubling time (lower figure) for the Lorenz system.

In the present work, the lead-time is used as the measurement for the predictability of the system. Once we have defined an index, we can study the local predictability of the Lorenz system according to its initial conditions. Figure 6 shows the predictability time structure obtained in the phase space. It can be observed that the predictability has a structure. Consequently, we can assign a predictability of the state just knowing the initial conditions of the system.

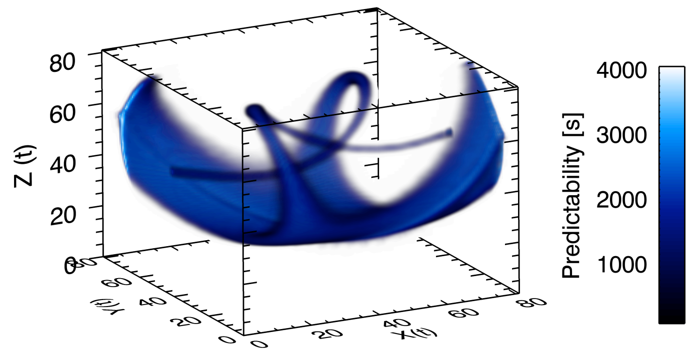

Figure 6.-3D predictability map in the phase space for the Lorenz system. The predictability is measure as the lead-time.

5. RAINFALL SYSTEM

Once, the main aspects of a well-known chaotic system as Lorenz one have been revisited. The main goal of this work is to realize an analog study for the rainfall fields. From our data set, it's obtained a 512x512 pixel field. Even though, we define two states for each pixel (rain/no-rain), the phase space obtained for the combination of all these pixels is of the order of $5 \cdot 10^{20}$. In our data set, there are 420'480 rainfall images. This means only a ~ 0.001% can be filled of the original phase space. Consequently, a new phase space has to be constructed by minimizing the number of variable to describe a rainfall field. The construction of the phase space is explained in subsection 5.1. After a lower dimensional phase space has been constructed, the attractor of the rainfall fields is classified (subsection 5.2). The inherent predictability and the local predictability of this attractor is studied in the last subsection (5.3).

5.1 *Phase space construction*

In order to reduce the phase space, rainfall fields are characterized by different statistics variables. The statistic parameters used in the present work are:

- **Marginal Mean:** Arithmetic mean of the rainfall pixels.
- **Standard Deviation:** Standard deviation of the rainfall pixels.
- **Skewness:** Third standardized moment of the rainfall pixels.
- **Kurtosis:** Fourth standardized moment of the rainfall pixels.
- **Coverage:** Percentage of rainfall pixels in function of total number of pixels.

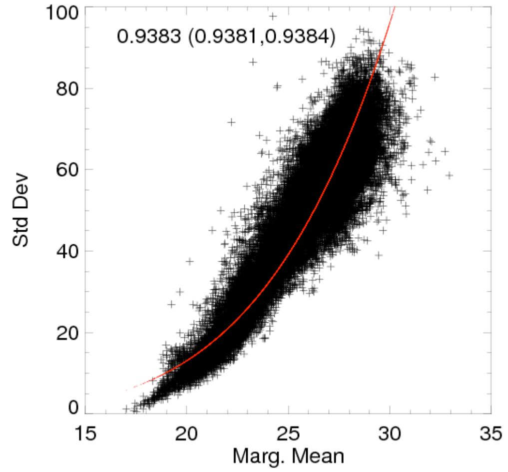

Figure 7.-Scatter plot of Marginal mean and Standard deviations. The Pearson correlation is computed by the red fitted line.

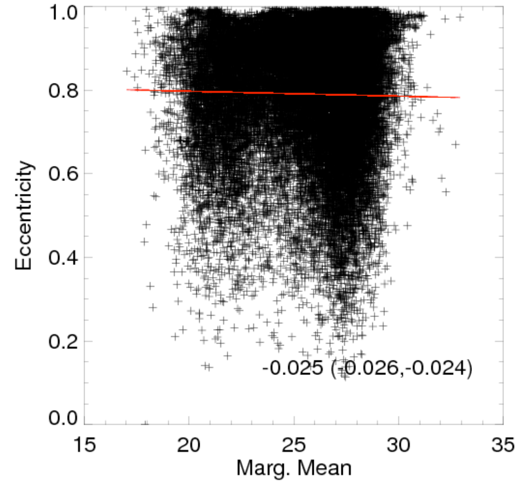

Figure 8.-Scatter plot of Marginal mean and Eccentricity. The Pearson correlation is computed by the red fitted line.

- **Number of cells:** Number of closed cells with an area bigger than 100 km2.
- **Power-spectrum slope:** Slope of the radially-averaged power spectrum.
- **Decorrelation distance:** Distance at which the autocorrelation function has 1/e value.
- **Eccentricity:** Eccentricity of the ellipse fitted over contour of 1/e. in the spatial autocorrelation field.
- **Orientation:** Angle of the same ellipse.

An important characteristic of the phase-space variables is that they are not correlated. The statistical parameters used to describe the rainfall field have been plotted in function of each other to analize the independence of them. It can be observed that some statistical parameters are highly correlated (Fig. 7). On the other hand, few statistical parameters are not correlated (Fig. 8). To define the phase space variables, just the uncorrelated variables are kept, whereas the correlated ones are removed. In the end, only three statistical parameters are used to define the rainfall images phase space: The marginal mean, the eccentricity and the decorrelation distance (or Area).

5.2 *Rainfall fields attractor*

Once the phase space variables are defined, the rainfall fields' attractor can be studied. As defined previously, the density gives information about the attractor for an experimental system as rainfall field where the exact equations are not formulated.

In Figure 9, the rainfall attractor is plotted. The first feature observed is the lack of a clear structure, as the butterfly wing of the Lorenz system had. It could be due to the scarcity of data as far as the data used is the time series instead of different realizations from perturbed initial conditions, but the density of points is quite similar to the obtained with the Lorenz system. It can be, most likely, the noise in the small scales caused by the fact that the reflectivity fields have some errors when measuring the rainfall (Zawadzki, 1975).

The correlation dimension is used to evaluate the self-similarity of this attractor. In figure 10, it can be observed the effect of the noise in the small scales (the scale region begins for larger radius than in the Lorenz system, Fig. 3). The correlation dimension obtained is 1.94. It is a non-integer number, which means that the simplification of the complex rainfall system to a phase space of three dimensions still keeps this system as a chaotic system.

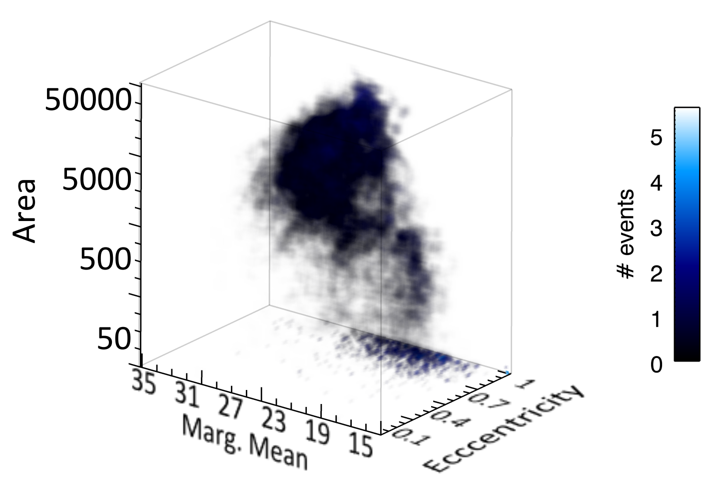

Figure 9.-3D plot of the density of points in the three statistical properties phase space for the rainfall fields. The numbers of events are per thousands.

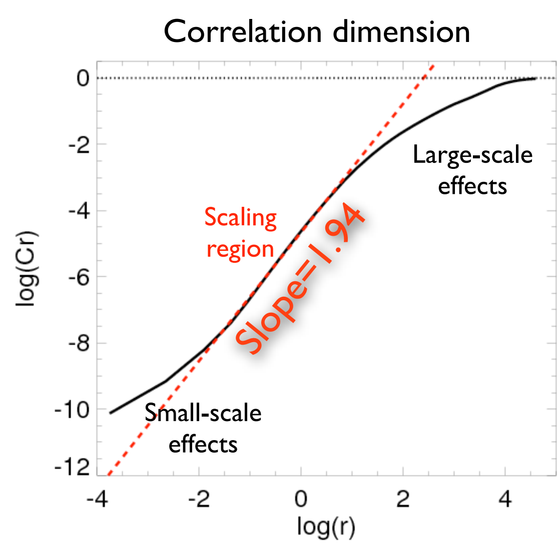

Figure 10.-The log(r)-log(Cr) plot for determining the correlation dimension by means of the scaling region slope.

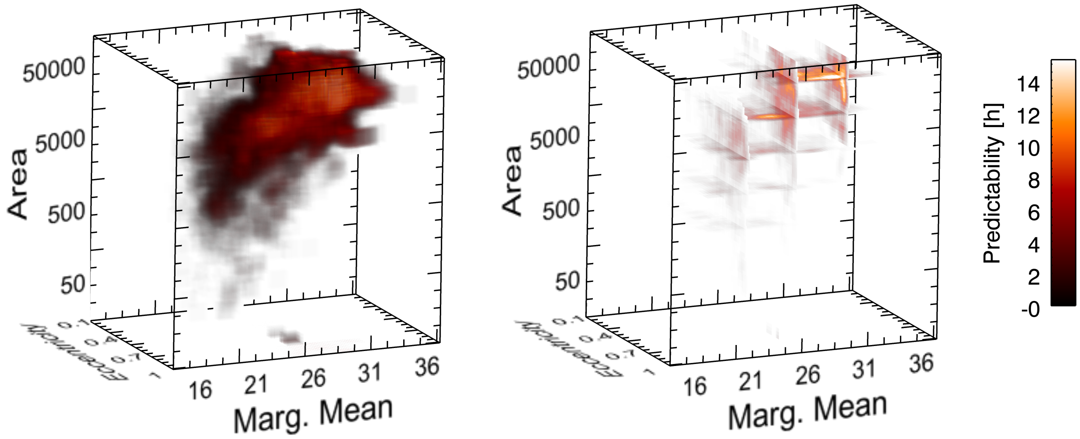

Figure 11.-3D predictability map in the phase space for the rainfall fields. The left side depicts the volume and the right side several planes are plotted in their actual position. Predictability is measure as the lead-time.

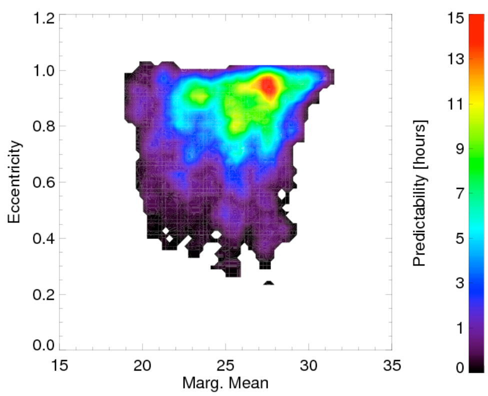

Figure 12.-2D projection of the predictability map over the marginal mean-eccentricity plane.

The fact that the correlation dimension is smaller than 2 is caused by the small, but still existing, correlation between the decorrelation distance and the marginal mean. This correlation reduces the degrees of freedom of the system. However, the fractal behavior of the system is clear and well fitted for the scaling region.

Finally, it has to be mentioned the large-scales effect. The scaling region is narrower than the obtained with the Lorenz system. The effect of the small scales is caused by the measurements errors in the rainfall fields. The large-scale effect could be caused by the clustered behavior of rainfall fields around the large mean and correlation distance area. In other words, the statistical properties are less variable for these kinds of precipitation systems than for the others.

5.3 *Predictability of the rainfall fields*

The predictability of rainfall fields can be measure by different indexes. In this work, as was discussed in section 4. The lead-time (time which the saturation is achieved) is used to characterize the predictability of the system. Figure 11 shows the 3D predictability map of the rainfall fields in the three statistical parameters phase space.

Even though, the boundaries of the surface of predictability are not well defined. It can be observed (in the right plot) that the interior part of the volume has a much more clear structure. In the 2D projection over the marginal mean-eccentricity plane (Fig. 12), it is clear the inside structure of the volume.

From this predictability map, it can be concluded that just from three statistical parameters of the rainfall field, the predictability of the system can be assessed. Besides, it is proven the fact that frontal systems are more predictable than MCS systems and these systems are, at the same time, more predictable than Isolated storms (a well-known fact in meteorology but that have not been tested from a chaotic point of view).

At this point, the predictability map tells us that the forecast lead-time depends only on some statistical parameters of the initial rainfall field. This information can be used to verified the add value or skill of a forecasting system. For this purpose, the variance of this map has to be studied. Figure 13 shows the standard deviation (square of the

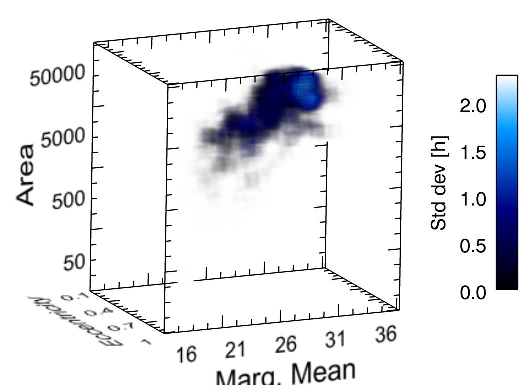

Figure 13.-3D representation of the standard deviation of the predictability time.

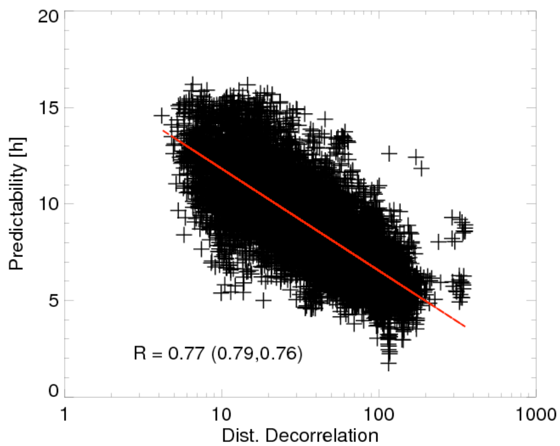

Figure 14.- Prediction time in function of the decorrelation distance. The Pearson correlation is computed by the red fitted line.

variance) of the predictability map. It can be observed that around 2 times the standard deviation will explain around the 95% of the variance. Consequently, in the area of longest predictability time an error of ±3 hours will explains a 95% of the variance of it.

This feature can be observed also looking into the relation between the statistical variables and predictability. Figure 14 shows an example of high correlated (R^2 = 0.77) relation between the decorrelation distance and the lead-time. The total variance explained is more than 80% with an error of around ±3 (as obtained by the standard deviation in the predictability map).

6. CONCLUSIONS

In this study, 15 years of rainfall fields at 15-minute temporal resolution have been used to study the predictability from a chaos theory point of view. The main conclusions can be summarized as:

- Three uncorrelated statistical properties of the rainfall fields have been chosen to construct the phase space.
- Rainfall fields in this phase space have a strange attractor with fractal structure and correlation dimension of 1.94.
- The rainfall field system in this new defined phase space is a chaotic system.
- A clearer interior structure can be observed in the predictability map.
- Inherent predictability can be determined by the initial statistical properties of the rainfall field.
- A ±3 hour error around the prediction time (computed by lead-time measurement) explains a 95% of the variance.

7. REFERENCES

Carbone, R.E., and J.D. Tuttle, 2008: Rainfall occurrence in the US warm season: the diurnal cycle. *J. Clim*. 21 (16), 4132–4146.

Grassberger, P. 1986: Do climatic attractors exist?. *Nature*, 323, 609-611.

Jankov, I., and W. A. Gallus, 2004: MCS Rainfall Forecast Accuracy as a Function of Large-Scale Forcing. *Wea. Forecasting*, 19, 428-439.

Keil, C., F. Heinlein, and G. C. Craig, 2013: The convective adjustment time-scale as indicator of predictability of convective precipitation. *Quarterly Journal of the Royal Meteorological Society*, n/a-n/a.

Lorenz, E., 1969: Atmospheric predictability as revealed by naturally occurring analogues. *Journal of the Atmospheric sciences*, 26 (4), 636–646.

Melhauser C. and F. Zhang, 2012: Practical and Intrinsic Predictability of Severe and Convective Weather at the Mesoscales. *Journal of the Atmospheric sciences*, 69, 3350-3371.

Nicolis, C. and G. Nicolis, 1984: Is there a climatic attractor?. *Nature*, 311, 529-532.

Smith, L. A., C. Ziehmann and K. Fraedrich, 1999: Uncertainty dynamics and predictability in chaotic systems. *Quarterly Journal of the Royal Meteorological Society*, 125, 2855-2886.

Toth Z. and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bulletin of the American Meteorological Society*, 74(12), 2317-2330.

Zawadzki, I., 1975: On radar-raingage comparison. *J. Appl. Meteor.*, 8, 1430-1436.